

## KORELACJA MIĘDZY OBIEKTAMI GRAFICZNYMI JAKO ASPEKT WYSZUKIWANIA ICH W OBRAZOWEJ BAZIE DANYCH

Tatiana Jaworska

*e-mail: [Tatiana.Jaworska@ibspan.waw.pl](mailto:Tatiana.Jaworska@ibspan.waw.pl)*

*Institut Badań Systemowych, Polska Akademia Nauk, ul. Newelska 6, 01-447 Warszawa*

**Słowa kluczowe:** przetwarzanie obrazu, obrazowa baza danych, klasyfikacja obiektów, zbiory rozmyte

### STRESZCZENIE

W miarę jak przybywa różnego rodzaju obrazów coraz istotniejsze staje się ich wyszukiwanie w bazach danych. W tym celu ważną rolę odgrywa kategoryzacja treści obrazów i wyszukiwanie konkretnych obiektów na obrazach. W budowanym przez nas systemie obrazy są segmentowane w celu wydzielenia obiektów. Segmentacja zachodzi na podstawie koloru i tekstury. Następnie obiekty są sklasyfikowane na podstawie wektora cech, do klasy określonej w bibliotece wzorców.

Informacja na obrazach jest indeksowana dwustopniowo. Na niższym poziomie obiekty są identyfikowane na podstawie wektora cech, który dla każdego obiektu jest porównywany z wektorem wzorcowym. Natomiast na wyższym poziomie zapisana jest informacja o rozmieszczeniu już sklasyfikowanych obiektów względem siebie jak i w całym obrazie. Dopasowanie obrazów następuje poprzez dopasowanie rozkładu przestrzennego obiektów w całym obrazie oraz poprzez obliczenie stopnia dopasowania poszczególnych obiektów do siebie.

Przy wyszukiwaniu odpowiedzi na zapytanie graficzne należy uwzględnić niedokładności wynikające z nieprecyzyjnej lokalizacji obiektów przez użytkownika, oraz skalowania tych obiektów w odpowiednim zakresie. Nieprecyzyjności te są uwzględniane w systemie poprzez zastosowanie trapezoidalnych zbiorów rozmytych. W rezultacie w odpowiedzi na nasze zapytanie obrazowe, otrzymujemy większą elastyczność wyszukiwania. Między innymi wyszukiwane są obrazy, które mogą być fragmentami lub większą częścią obrazu stanowiącego zapytanie lub jego lustrzanym odbiciem.

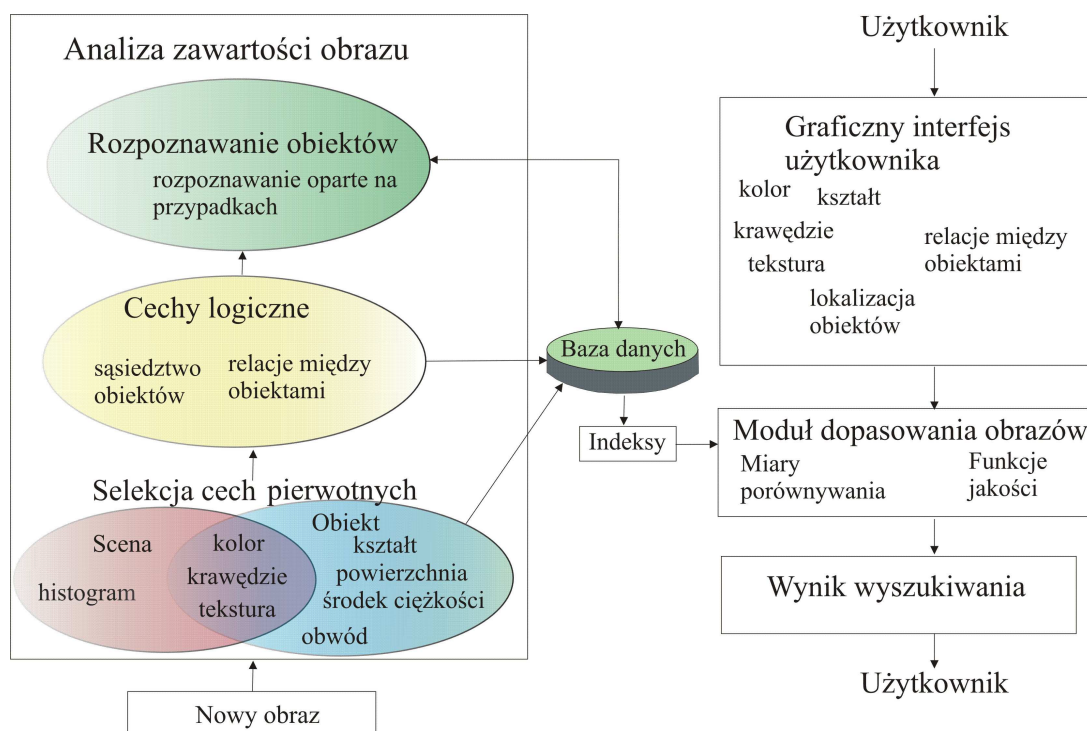
### 1. Informacje ogólne o obrazowej bazie danych

Budowa obrazowej bazy danych wymaga implementacji takich operacji jak: analiza zawartości obrazu, indeksowanie, przeszukiwanie, porównywanie, formułowania zapytań itp., ale przy użyciu mechanizmów wewnętrznych umożliwiających wykonywanie tych wszystkich operacji na obrazach [1]. Trzeba przy tym uwzględnić wiele rodzajów niepewności lub wręcz braku informacji takich jak: zaszumienie lub zniekształcenie obrazu, nieprecyzyjne zapytanie konstruowane przez użytkownika, błędy w segmentacji obrazu wynikające z różnego oświetlenia.

W budowanym przez nas systemie [2] wyszukiwania obrazu są segmentowane w celu wydzielenia obiektów. Nowy obraz jest poddawany obróbce wstępnej polegającej na wyekstrahowaniu z niego cech, tak dotyczących całego obrazu, jak i wyszczególnieniu obiektów w obrazie. Obiekty graficzne są opisywane poprzez cechy niskiego poziomu takie jak: średni kolor, środek ciężkości, powierzchnia, obrys, parametry tekstury, mimośród, orientacja, itp. Na ich podstawie jest tworzony wektor cech. W dalszej kolejności obrazy i wszystkie policzone wcześniej cechy są wprowadzane do bazy danych. Następnie obiekty są identyfikowane poprzez porównanie ich wektora cech z wyznaczonym wzorcem i są etykietowane w sposób zrozumiały dla użytkownika np. dom, okno, drzewo, itp.

Informacje o całych obrazach, wydzielonych obiektach, jak i biblioteka wzorców dla obiektów są przechowywane w bazie danych w systemie Oracle ver. 10g. Oprogramowanie dotyczące przetwarzania obrazu jest wykonane w MatLabie. ver. 2010a.

Cechy obiektów graficznych są reprezentowane na trzech poziomach poczynając od informacji o kolorze, teksturze i deskryptorach kształtu dla poszczególnych obiektów. Następnie ustalane są wzajemne relacje przestrzenne między obiektami. Na najwyższym poziomie znajduje się moduł analizujący obraz w kategoriach ludzkiej percepcji. Hierarchia ta jest widoczna w lewym bloku rys. 1.



Rysunek 1 Ogólna struktura obrazowej bazy danych

## 2. Schemat indeksowania obrazów

Cechą globalną jest przestrzenne rozmieszczenie obiektów graficznych w obrazie. W tym wypadku interesuje nas rozmieszczenie środków ciężkości obiektów  $(x_c, y_c)$  i ich identyfikatory  $o_N$ . Dla wszystkich obiektów w obrazie tworzymy wektor  $F_g$ , gdzie  $N$ -liczba obiektów w obrazie:

$$F_g = \begin{bmatrix} (x_{c_1}, y_{c_1}), o_1 \\ (x_{c_2}, y_{c_2}), o_2 \\ \vdots \\ (x_{c_N}, y_{c_N}), o_N \end{bmatrix} \quad (1)$$

Każdy obiekt opisujemy przy pomocy zbioru cech takich jak: średni kolor  $k_{av}$ , parametry tekstury  $T_p$ , powierzchnia  $A$ , powierzchnia wielościanu wypukłego  $A_c$ , współrzędne środka ciężkości  $\{x_c, y_c\}$ , mimośród  $e$ , moment bezwładności i momenty Zernike'a, osie symetrii  $m_{long}$  i  $m_{short}$ , etc.

Na podstawie tego zbioru cech  $F = \{k_{av}, T_p, A, A_c, \dots, E\}$  w zapisie uproszczonym  $F = \{f_1, f_2, \dots, f_r\}$ , gdzie  $r$ -liczba cech, przyjmujemy wektor cech  $O$  do porównania obiektów z wzorcami cech [3].

$$O = \begin{bmatrix} O(k_{av}) \\ O(T_p) \\ \vdots \\ O(E) \end{bmatrix} = \begin{bmatrix} O(f_1) \\ O(f_2) \\ \vdots \\ O(f_r) \end{bmatrix} \quad (2)$$

## 3. Biblioteka wzorców

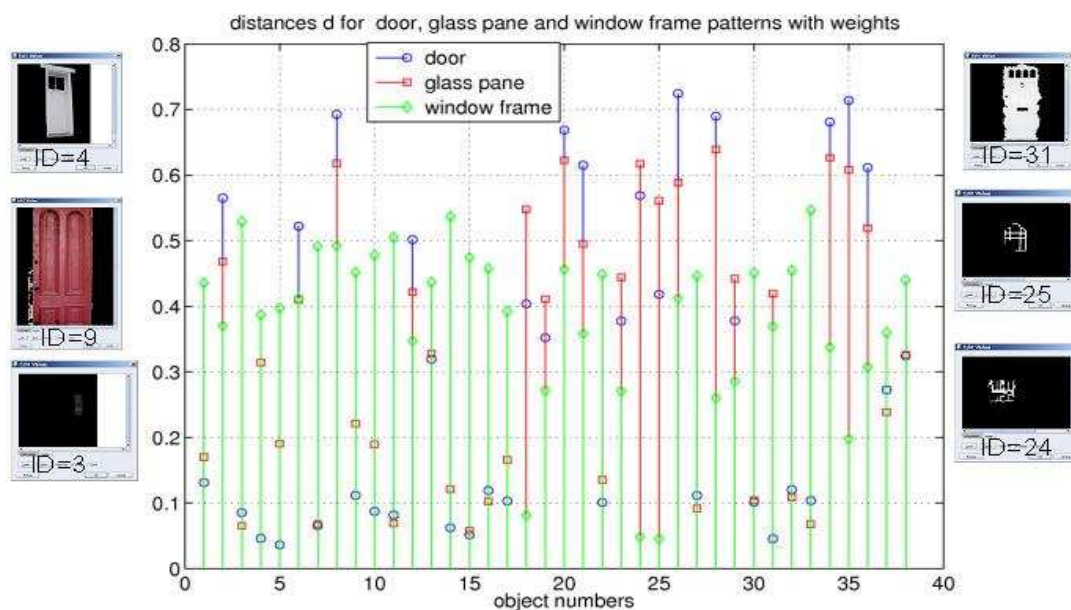
Biblioteka wzorców zawiera informację o typie wzorca, lokalizacji obiektu i zakresie wartości charakterystycznych dla danego obiektu graficznego. Dla każdego obiektu określamy wzorcowy wektor cech  $P_k$  i przyjmujemy wagi  $\mu_{P_k}(f_i) \in [0,1]$ .

Każdy obiekt jest przyporządkowywany do określonej klasy z biblioteki wzorców poprzez szukanie najmniejszej odległości pomiędzy wektorami  $O$  i  $P_k$

$$d(O, P_k) = \left[ \sum_{i=1}^r \mu_{P_k}(f_i) |O(f_i) - P_k(f_i)|^m \right]^{1/m} \quad (3)$$

gdzie:  $m=1,2$ . Na tej podstawie każdy obiekt otrzymuje etykietę  $o_k$ .

Rys. 2 przedstawia odległości  $d$  czyli stopień zgodności cech obiektów z 3 przykładowymi wzorcami (drzwi, szyby i framugi okienne). Jako próbkę do klasyfikacji przyjęto 38 obiektów.



Rysunek 2 Klasyfikacja obiektów na podstawie wzorców

#### 4. Informacja niepewna w obrazowej bazie danych

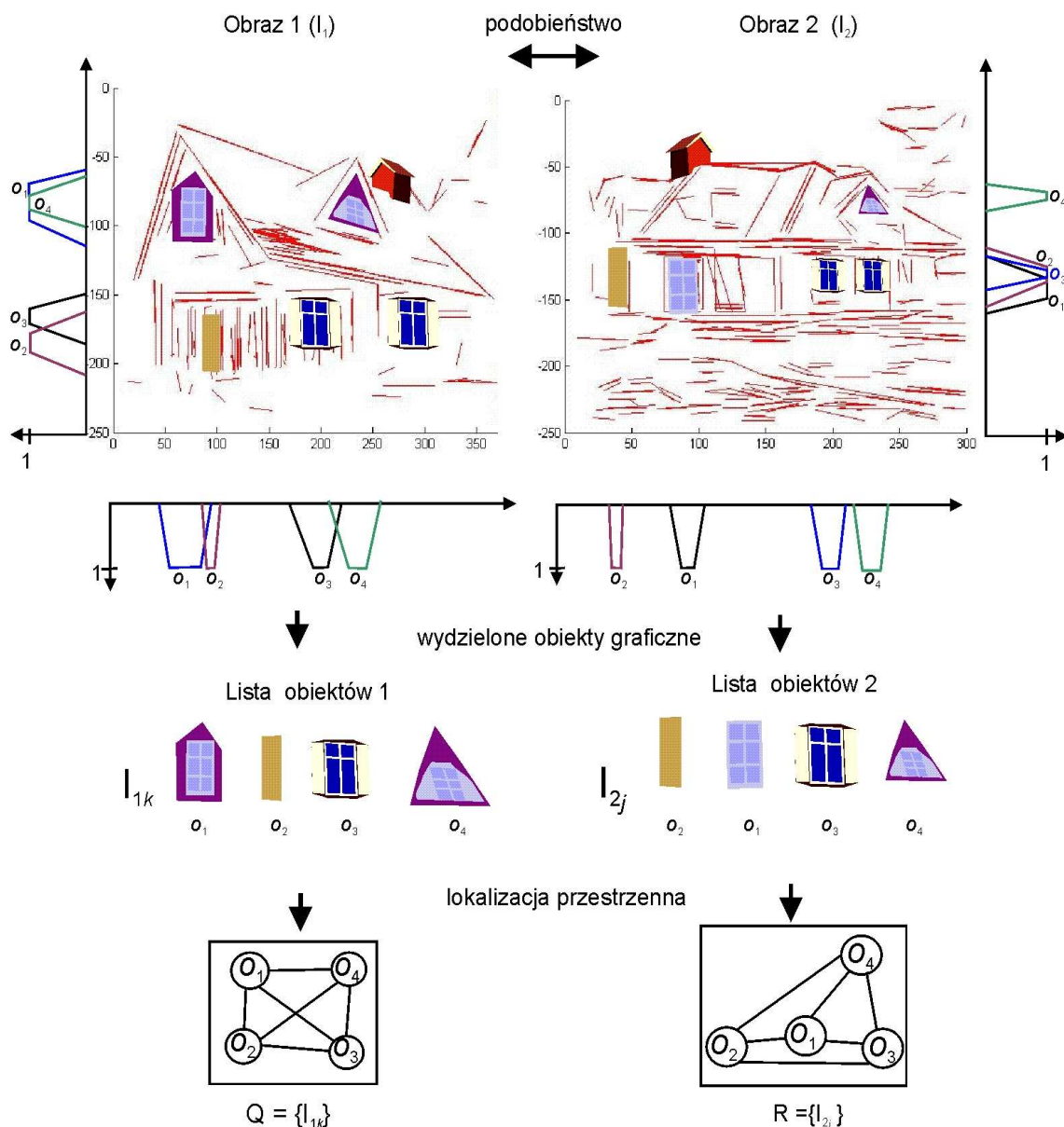
Rys. 3 prezentuje w sposób ideowy metodę porównywania obrazu  $I_1$ , który jest zapytaniem użytkownika, z obrazem  $I_2$  przechowywanym w bazie danych. Porównywanymi elementami są: występujące obiekty  $o_1, o_2$ , itp. oraz stopień ich wzajemnego podobieństwa na obu obrazach oraz ich lokalizacja przestrzenna tak w całym obrazie jak i wzajemnie w stosunku do siebie.

W naszym systemie użytkownik może skonstruować własne zapytanie graficzne korzystając z istniejących już obiektów graficznych. Obiekty te można organizować przestrzennie na szkicach obrazów. W tym celu wykorzystuje graficzny interfejs użytkownika GUI (Rys. 1).

Dla każdego obiektu przyjmujemy, że został on umieszczony na obrazie przez użytkownika z pewną niedokładnością, co można zapisać jako wektor  $F_g$ .

$$F_g = \begin{bmatrix} \min(\mu(x_{c_1}), \mu(y_{c_1})), o_1 \\ \min(\mu(x_{c_2}), \mu(y_{c_2})), o_2 \\ \vdots \\ \min(\mu(x_{c_N}), \mu(y_{c_N})), o_N \end{bmatrix} \quad (4)$$

gdzie:  $N$  oznacza liczbę obiektów w obrazie,  $\{x_{c_i}, y_{c_i}\}$  - współrzędne środków ciężkości obiektów,  $\mu(x_{c_i}), \mu(y_{c_i})$  - trapezoidalne funkcje przynależności opisująca niedokładność lokalizacji środków ciężkości  $x$  i  $y$  obiektów przez użytkownika,  $\min(\mu(x_{c_i}), \mu(y_{c_i}))$  - operator porównania funkcji rozmytych,  $o_k$  - etykieta obiektu przypisywana w procesie identyfikacji. Dodatkowo przyjmuje się też niedokładność skali dla poszczególnych obiektów umieszczanych przez użytkownika.



Rysunek 3 Ogólny schemat szukania podobieństwa pomiędzy dwoma obrazami wewnątrz obrazowej bazy danych. Po identyfikacji obiektów  $o_i$ , sprawdza się ich rozmieszczenie na obrazie. Nieprecyzyjność lokalizacji środka ciężkości opisana jest przy pomocy zbiorów rozmytych, co obrazują wykresy z trapezami.

Moduł dopasowania obrazu wybiera obrazy, które mają równą lub zbliżoną liczbę takich samych obiektów, a układ przestrzenny tych obiektów jest najbliższy zapytaniu użytkownika. Następnie moduł ten wysyła swoje wyniki do Oracle'a, w którym następuje wyszukanie i wyświetlenie zbioru obrazów najlepiej spełniających zapytanie użytkownika [4].

Tabele 1 i 2. Informacja przestrzenna dla obrazu  $I_1$  i dla odpowiadającego mu obrazu  $I_2$  z Rys. 2.

$I_1$	$o_1$	$o_2$	$o_3$	$o_4$	$I_2$	$o_1$	$o_2$	$o_3$	$o_4$
$o_1$	0	S	SE	E	$o_1$	0	W	E	SW
$o_2$	N	0	E	NE	$o_2$	E	0	E	NE
$o_3$	SW	W	0	S	$o_3$	W	W	0	SE
$o_4$	W	NW	N	0	$o_4$	NE	SW	NW	0

### 5. Momenty Zernike'a jako deskryptory kształtu obiektów

Poza wspomnianymi już cechami niskiego poziomu obiekty graficzne są opisywane przez deskryptory kształtu. Na ich podstawie tworzony jest wspomniany też wektor cech. Jako główne deskryptory zostały przyjęte momenty bezwładności

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad p, q = 0, 1, 2 \quad (5)$$

i momenty Zernike'a [5]. Zespolone momenty Zernike'a wyprowadzamy z ciągu wielomianów ortogonalnych nad dyskiem jednostkowym

$$V_{pq}(x, y) = V_{pq}(r, \theta) = R_{pq}(r) e^{jq\theta}, \quad r \in [-1, 1] \quad (6)$$

gdzie  $r = \sqrt{x^2 + y^2}$  jest długością wektora od środka dysku do punktu  $(x, y)$ , a  $R_{pq}(r)$  wielomiany o wartościach rzeczywistych:

$$R_{pq}(r) = \sum_{s=0}^{(p-q)/2} (-1)^s \frac{(p-s)!}{s! \left(\frac{p+|q|-s}{2}\right)! \left(\frac{p-|q|-s}{2}\right)!} r^{p-2s} \quad (7)$$

gdzie  $p - |q|$  jest parzyste,  $0 \leq |q| \leq p$  i  $p \geq 0$ .

Niech  $s \rightarrow (p - k)/2$ , wtedy wielomiany mogą być reprezentowane jako:

$$V_{pq}(x, y) = \sum_{k=q}^p B_{pqk} r^k e^{jq\theta} \quad (8)$$

Wtedy  $B_{pqk}$  są współczynnikami tego wielomianu:

$$B_{pqk} = \frac{(-1)^{(p-k)/2} \left(\frac{p+k}{2}\right)!}{\left(\frac{p-k}{2}\right)! \left(\frac{q+k}{2}\right)! \left(\frac{k-q}{2}\right)!} \quad (9)$$

Zatem, dwuwymiarowe momenty Zernike'a rzędu  $p$  z powtórzeniami  $q$  dla obrazu o funkcji intensywności  $f(x, y)$

$$Z_{pq} = \frac{p+1}{\pi} \iint_{x^2 + y^2 \leq 1} V_{pq}^*(x, y) f(x, y) dx dy \quad (10)$$

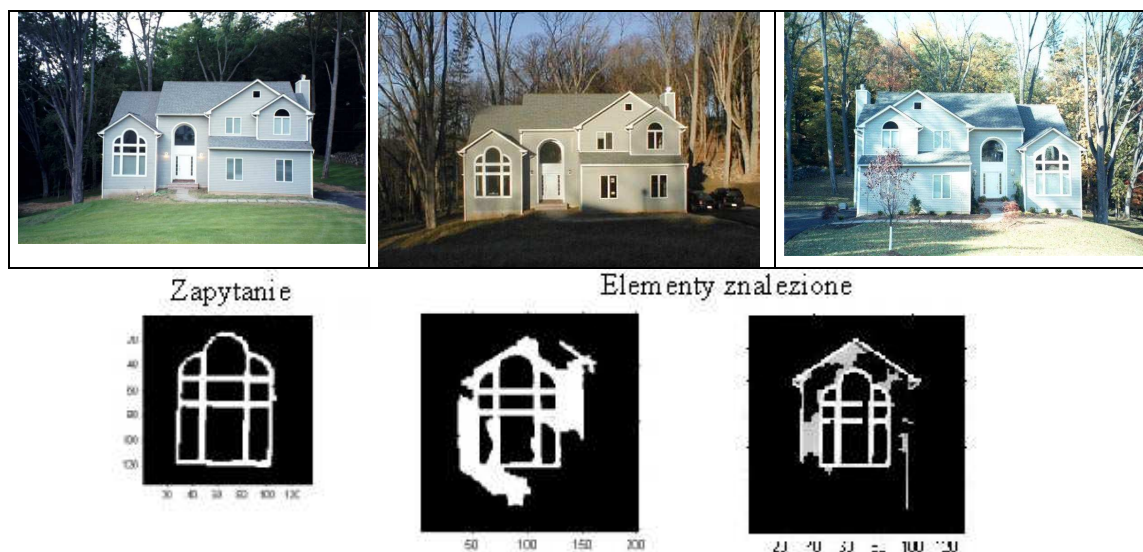
gdzie

$$V_{pq}^*(x, y) = V_{p, -q}(x, y) \quad (11)$$

Właściwości momentów Zernike'a:

1. Tak zdefiniowane momenty Zernike'a są niezmiennicze względem rotacji.
2. Aby uzyskać niezmienniczość względem przesunięcia trzeba centrum obiektu umieścić w środku układu współrzędnych.
3. Natomiast niezmienniczość skali uzyskujemy poprzez skalowanie obiektu.





Rysunek 4 Przykładowe okno znalezione na podstawie najmniejszej odległości euklidesowej dla momentów Zernike'a do rzędu trzeciego z podanych powyżej zdjęć budynku w różnych oświetleniach (prawe zdjęcie w odbiciu lustrzanym).

## 6. Wnioski i kierunki dalszych prac

Analiza kontekstowa obrazów jest w tej chwili jednym z najtrudniejszych zagadnień z jakim zmagają się przetwarzanie obrazów. W naszym instytucie podejmujemy próbę zbudowania takiego systemu w oparciu o możliwie szeroki wachlarz metod analitycznych. Obecnie rozbudowywany jest prototypowy system łączący możliwości obliczeniowe Matlaba z bazodanowymi zaletami systemu Oracle.

W najbliższej przyszłości będzie wprowadzana implementacja weryfikująca poprawność metod wyszukiwania w przypadku dużej różnorodności obrazów. W tym celu będzie też rozszerzana biblioteka wzorców o inne rodzaje obiektów graficznych. Tam gdzie będziemy mieli do czynienia z informacją niepewną i nieprecyzyjną szeroko będzie stosowana analiza oparta o zbiory rozmyte, przez co będziemy dążyć do zwiększenia uniwersalności tworzonego systemu.

## LITERATURA

- [1] Deb, S. (red.), "Multimedia Systems and Content-Based Image Retrieval", IDEA Group Publishing, Melbourne, (2004).
- [2] T. Jaworska, "Object extraction as a basic process for content-based image retrieval (CBIR) system." W: Opto-Electronics Review, Vol.15, Nr. 4, str. 184-195, (2007).
- [3] T. Jaworska, „Multi-criteria Object Indexing and a Graphical User Query as an Aspect of Content-based Image Retrieval System”, Information Systems Architecture and Technology, Advances in Web-Age Information Systems, str. 103-112, Wrocław, (2009).
- [4] T. Jaworska, J. Kacprzyk, N. Marin, S. Zadrozny, "On dealing with imprecise information in a content-based image retrieval system", Computational Intelligence for Knowledge-Based Systems Design, 13th International Conference on Information Processing and Management of Uncertainty, IPMU 2010, Dortmund, Niemcy, 28 czerwca – 2 lipca, Lecture Notes in Computer Science, Vol. 6178/2010, str. 149-158, (2010).
- [5] M.R. Teague, "Image analysis via the general theory of moments", JOSA, Vol. 70, Wyd: 8, str. 920-930 (1980)